

## One More Datum on Rorschach Form Quality

DANILO R. SILVA<sup>1</sup> AND ANTÓNIO A. PIRES<sup>2</sup>

<sup>1</sup>*Faculty of Psychology, University of Lisbon*

<sup>2</sup>*Faculty of Psychology, University of Porto*

The central purpose of this work is to examine to what extent Form Quality (FQ) scoring of Rorschach responses, using Exner's (2003) Table A, is the reason for lower FQ to often be found in normative data outside the United States. The Rorschach protocols of 180 Portuguese boys and girls between 6 and 10 years of age were codified with respect to Form Quality in accordance with Exner's table of objects classified as ordinary, unusual, and minus, based on frequencies seen among U.S. adults and on an analogous Portuguese table (Table P) constructed from the protocols of 400 nonpatient children. Mean differences between the various FQ variables of the 2 tables were found to be insignificant. However, use of Janson's (2003) iota coefficient to evaluate agreement between the 2 approaches to scoring on the same protocols revealed that the more striking differences occurred in the FQu variable when agreement was assessed on the basis of each individual response. In view of these results, the authors recognize the possibility of Rorschach normative data to be universally applied, but still consider the possible advantage of each country or culture to build its own FQ table that will permit preservation of its particularities.

The article published by Meyer, Erdberg, and Shaffer (2007) in the *Journal of Personality Assessment* Supplement on International Reference Samples for the Rorschach Comprehensive System appears to reopen the issue raised by the persistent discrepancies between normative data gathered in the United States and published by Exner (1985, 1990, 1995), and normative data published in the last 15 years in Europe and Latin America. In fact, following a paper presented by one of us at the 14th International Congress of Rorschach and Other Projective Methods, appropriately titled "Are Rorschach Normative Data Presented by Exner Valid in extenso for the European Population? An Essay With a Portuguese Sample" (Silva, Novo, & Prazeres, 1990), a number of new Rorschach normative data studies flourished in Europe, South America, and a little later in North America and Japan. All these studies made the same point, that in most of the studies referred to earlier, means of several variables relating to Form Quality (FQ) were too far below Exner's.

Given the referred to disparity within the new normative data, we were surprised to read the Meyer et al. (2007) article openly supporting the claims for universal Rorschach norms, already present in Exner and Weiner's (1995) writings, as in the following passage related to the development of FQ tables:

These findings might lead some to argue that normative data for the test should be established by country, or language, or even by culture. Such a suggestion is probably not very realistic unless very large samples are available. Instead, it seems more reasonable to suggest that the items in the Form Quality Table be reviewed for frequency. In other words, if a response currently not found in the Table occurs frequently in a given country or culture, the form quality scoring for that answer should be adjusted accordingly. Likewise, if a response that is currently listed in the Form Quality Table as unusual occurs with a high frequency in the nonpatient records for a country or culture, it should be scored as ordinary. This is a very different procedure from that involved in

attempting to establish separate normative data. (Exner & Weiner, 1995, p. 50)

It is clear, according to these authors, that American norms are appropriate for all countries and cultures, although they admitted that each one might introduce some changes according to the frequency of the respective answers. Let us consider, now, the following passage from the *Journal of Personality Assessment* supplement:

These findings indicate that, by and large, adults look pretty much the same on the CS [Comprehensive System] no matter what language they speak, what country they reside in, and what cultural background influences them. A question remains about whether the relatively small variation between countries is due to differences in culture, language, participant selection criteria, administration standards, coding benchmarks and/or examiner skill. Despite all these potential influences, because the between-sample influences are modest, the findings support the transportability of the Rorschach across countries. (Meyer et al., 2007, p. S206)

Here is the confirmation of Exner and Weiner's idea.

Notwithstanding the directness of Meyer et al.'s (2007) statements, some comments attempting to explain or justify differences between Exner's normative values and those that appeared in other countries deserve to be mentioned here. FQ is undoubtedly the variable that better expresses such differences, either because it is present in almost all responses as their determinant or because it better translates their perceptual basis. Ephraim (2000) and Andronikof-Sanglade (2000) presented two different approaches, both reflecting cultural standpoints, to justify normative differences between FQ means found by Exner and those found outside the United States. Ephraim (2000) pointed out:

The participants from the latter samples (Southern European and Latin American countries) seem to be comparatively more inclined to personalize the stimulus. In Piagetian terms, they tended to assimilate the stimulus to their internal reality rather than accommodate their responses to the inkblot perception demands. Conversely, Exner's subjects could

Received March 1, 2010; Revised December 19, 2010.

Address correspondence to Danilo R. Silva, Alameda Antonio Sergio, 1–3D, Lisbon 1750–031, Portugal; Email: danilo@fpce.ul.pt

be described as comparatively more inclined toward an instrumental, problem-solving, consensus-oriented approach to the Rorschach task, as expressed by their tendency to be more realistic and precise in their perceptions. (p. 313)

In other words, Ephraim justifies mean FQ differences in terms of the individual's different attitudes toward the task, which stem from different cultural bases.

On the other hand, Andronikof-Sanglade examines the FQ problem from a European perspective, arguing that it has historically taken, and should continue to take into account the similarity between the form of the object indicated in the answer and the form of the blot or of the blot area chosen by the subject, regardless of the frequency of the answer. She attempts to explain the wide discrepancies ascertained between Exner's normative values for X+% and the much lower European normative values. An important reason for this disparity comes from the fact that, according to Andronikof-Sanglade (2000), Exner's definition of good form takes as its main criterion "the frequency of occurrence of a specific object as designated by a word as opposed to a frequent occurrence of a specific shape identifying a certain class of objects meeting the form requirements of the blot" (pp. 337–338). The author underlines the fact that Exner emphasizes the frequency of the response in the sample against the appropriateness of form of the object to the corresponding blot area, the chief criterion claimed to be used by all the other authors, particularly the European ones, in coding FQ. A second important detail to which Andronikof-Sanglade calls our attention is related to the role of language. She writes:

By content we mean the words to designate an object or shape. In that sense, content is a product of language, and we know that language is the making of culture. We are born in a culture, and we encounter the world through a language. Language is the utmost expression of a culture, its identity mark, both its core experience and symbol. (p. 336)

From this statement the author goes on to identify the Rorschach Content category as the one most affected by culture. So the act of scoring protocols, European protocols, following Exner's FQ table, implies lower X+% mean values, a fact that "can only be interpreted as evidence of language and environmental differences across cultures." And the author adds, "Until proper lists of frequent responses are established for the European countries, no interpretation in terms of socially acceptable behaviors, individualistic orientation, or reality testing would be legitimate" (Andronikof-Sanglade, 2000, p. 338).

From this perspective, cultural differences, particularly those transmitted by language, appear to be the reason for the differences found between American and European FQ means. As a conclusion of her overview of the Comprehensive System in Europe and of the available data, Andronikof-Sanglade (2000) considered "that extrapolation of the RCS [Rorschach Comprehensive System] to non-American cultures requires some adjustment and cannot be directly made by transferring normative data" (p. 343).

Based on these considerations, we should look again at some of the aforementioned studies undertaken and developed in the United States that are related to Exner's normative data. At the end of the 1990s, some published papers, such as those of Erdberg and Shaffer (1999) or Shaffer, Erdberg, and Haroian (1999) showed results that departed from Exner's on FQ, the crucial variable of interest here. That fact is well illustrated in

a review by Viglione and Hilsenroth (2001), who showed the mean X+% was lower and the mean X-% and Xu% was higher in U.S. nonpatient samples when compared to Comprehensive System normative values. Exner's (2005) own new normative data show similar changes in the mean of FQ values. Also Meyer (2001), in an excellent comment on Wood, Nezworski, Garb, and Lilienfeld's (2001) strong criticism of the Rorschach and the Comprehensive System, pointed out that "each time an IQ test is renormed, the same level of cognitive capacity is associated with a lower score" (p. 389), a fact about cultural change that might justify finding lower scores on X+% and higher scores on Xu%, X-%, S-%, and M-% across time (Meyer, 2001).

The publication of the *Journal of Personality Assessment* supplement (Shaffer, Erdberg, & Meyer, 2007) not only was a very strong contribution to the understanding of the problem of FQ differences that had emerged outside of the United States, but also affected perceptions in the United States on the matter. Results obtained among the contributing countries, from east to west, provided the researchers with new sources of information and the ability to compare their results with others, engendering both thoughtful criticism and positive feelings about their similarities on FQ. On the other hand Meyer et al.'s (2007) proposal toward the establishment of international Rorschach norms, on the basis of the absence of significant differences, both between the normative data for each country, and between the whole international sample and Exner's (2001) CS normative data, was an important advance that astonished the majority of researchers, who, in spite of their different reactions, took the line it would be good to reflect on this possibility. It should be noted, however, that for evaluated data to be correctly referred to, we take note of Meyer and Viglione's (2008) point concerning the comparisons between the values of the international sample and those of the 25-year-old Exner sample (Exner, 2001):

People in the composite international sample used more unusual location areas, incorporated more white space, had less healthy form quality scores, made less use of color, tended to see more partial rather than full human images, and showed a bit more disorganization in thinking. (p. 319)

All these sources bring us to a better understanding of the obtained data and signal an important advance in scientific knowledge.

Sharing these concerns, we endeavored to proceed with our studies of a sample of children. Within the scope of the international study described in the previously mentioned *Journal of Personality Assessment* supplement, results from child samples were revealed to be inconsistent when compared with those from adult samples. The following quotation throws light on the origins of the inconsistency:

The child and adolescent data are incomplete, emerging from just five countries that also varied in their sampling of the full spectrum of development and their grouping of samples across ages. This makes it difficult to disentangle what may be influences due to culture, age, or administration and scoring effects. (Meyer et al., 2007, p. S211)

This is an additional reason to persist in our study. Because Exner's Table A, the instrument for assigning FQ coding to answers on the Rorschach, is based on protocols that do not include protocols of children, we created a table to establish Rorschach normative FQ data for 6- to 10-year-old Portuguese

children. Specifically, an FQ table was created from a set of 400 protocols of children who regularly attended one school in Lisbon and various state schools on the south bank of the Tagus. In doing so we sought to obtain, as far as the FQ is concerned, genuine Portuguese FQ values entirely independent of the American coding source.

### METHOD

The sample is made up of 40 children of both genders aged 6, 7, 9, and 10, and 20 children aged 8. Most of these children attended state schools within the Lisbon metropolitan area and came from middle-class and working-class social backgrounds. A small number came from the middle to upper social levels, attending private schools. All the children were examined with their own consent and that of their parents. They were nonpatients and attending the normal school years corresponding to their age. Statements from their teachers attested that they were suffering from no learning difficulties. Children were tested in their schools. All examiners had at least two semesters on the Comprehensive System and two other semesters on coding, administration, and interpretation training.

These 180 protocols were randomly extracted from the Portuguese sample of 357 children from which normative data were established (Silva & Dias, 2007b) and who were part of the whole set of 400 children aged between 6 and 10 who were the original subjects contributing to the Portuguese FQ table. In this sense, our sample is not an independent sample in terms of provenance, and that aspect contributes, in principle, to obtaining higher average FQ scores for our sample than for an independent sample. In effect, as an FQ measuring instrument, whereas Table A is totally independent of the protocols considered here, Table P is built in part from the same protocols that are being measured here. The 180 protocols coded for FQ according to Exner's Table A (2001, pp. 107–167) were recoded according to the new Portuguese Table, hereafter referred to as Table P.

We applied this new instrument, the first created outside the United States, in accordance with the criteria set by Exner to build his table. It should be noted that recoding was carried out by the same coders as the ones who coded our normative sample of 357 children (Silva & Dias, 2007b). This set of 180 protocols is used in this study to assess the effect of each table on the FQ mean values.

Exner's criteria for the construction of his table are described next, along with the specific criteria that were used for our table. FQ criteria are conditional on the location selected for the response, so locations are described first. Location use is defined by four scores. A whole (W) is scored for answers that use the whole blot. A common detail (D) is scored for answers that use part of the whole blot, with that area chosen for use by at least 5% of the total, which in this case is 20 children. An unusual detail (Dd) is scored for answers that use an area of the whole blot that is chosen by less than 5% of the total (i.e., less than 20 children in our sample). Finally, space (S) is coded for answers that use or are located in the noninked or blank areas of the inkblot.

Exner's FQ table differentiates four types of codes. A common or ordinary response (FQo) to a W or D location is defined as an object seen by at least 2% of the total sample in a blot area with contours that actually exist. In this case, this corre-

sponds to objects seen by at least 8 children. In the case of a Dd location, FQo was scored for objects identified by at least two thirds of the subjects who chose the location together with blot contours that really exist. An uncommon response (FQu) is one that does not meet the criteria for FQo but the form of the object is, nevertheless, accepted as a good form by a jury of three qualified members. A distorted or minus response (FQ–) is scored when the form of the object is incompatible with the formal structure of the blot area being used and thus constitutes ipso facto a perceptual distortion or violates some law of visual perception, as often occurs in the case of the closure law, a principle according to which the subject looking at a figure with incomplete lines tends to complete (close) it and perceive it as complete. Finally, the fourth FQ type (FQ+) corresponds to an "ordinary-elaborated" form, one whose articulation is unusually enriched by the details the subject uses in his or her description, without reducing its appropriateness (Exner, 2003, pp. 122–123). It must be emphasized that the decision of scoring FQu for a concrete object whose FQ could not be scored was the result of independent judgments by three qualified members. These scores correspond to responses unanimously considered as easily and quickly seen and adjusted to the contours of the area. Whenever all three judges agreed on the FQ of the object, it was scored as u; disagreement between judges implies the use of FQ– score.

Interjudge reliability concerning the use of Table P for recoding FQ on the 180 Rorschach protocols was calculated using Janson's iota statistic and his Rorschach Utilities Program (Janson, 2003). Thirty-five protocols were independently recoded by both authors for FQ. Table 1 presents results concerning FQo, FQ–, and FQu in terms of percentage of agreement and iota at the response level, and iota at the protocol level. They are fairly acceptable and permit us to exclude the idea that results to be presented might be the result of deficiencies when coding FQ according to Table P.

### Statistical Analysis

With respect to mean differences, the following five protocol-level Comprehensive System variables that indicate FQ were considered: X+%, X–%, Xu%, WDA%, and XA%. Means and standard deviations of each variable were calculated and the significance of differences between means was evaluated by paired sample *t* test for the equality of means. As an index of effect size, Cohen's *d* was computed from the means and pooled standard deviations. A power analysis was completed using G-Power (Faul, Erdfelder, Lang, & Buchner, 2007). For a paired sample *t* test with 180 protocols and alpha set at .05, this study has 80% certainty (i.e., power = .80) to detect a small effect size ( $d = 0.21$ ) if it existed in the population.

TABLE 1.—Values of iota for interrater reliability when coding form quality at the response and the protocol level in a random sample of 35 of the 180 Portuguese child Rorschach protocols.

Variables	Response Level % of Agreement	Response Level Iota (Kappa Equivalent)	Protocol Level Iota (Weighted Kappa Equivalent)
FQo	.96	.92	.95
FQ–	.97	.93	.96
FQu	.93	.82	.84

To objectively evaluate the effect of applying one table or the other when coding the FQ of the responses, we used Janson's iota statistic and his Rorschach Research Utilities program (Janson, 2003). Iota, which is a multivariate extension of Cohen's kappa (Janson), is generally used to assess interjudge reliability (i.e., two judges coding the same protocols), as was done earlier. However, in fact, the situation is analogous for these comparative data because we have two different FQ coding tables applied to the same protocols. In this perspective, as in the case of interjudge reliability, more frequent differences between the two encodings are expected to yield lower iota values; similar classifications with the use of two different tables is expected to yield higher iota values. For these analyses we examined agreement results for eight variables at the protocol level (FQo, FQ—, FQu, X+%, X—%, Xu%, WDA%, and XA%), three of which were also considered at the response level. The response level reliability is of particular interest in our study because it permits us to see to what extent both Tables A and P code equally or differently the same single response.

### RESULTS

Table 2 shows means and standard deviations for each FQ variable. No significant differences were found between means of the American and Portuguese tables for the five variables, with effect sizes from near zero up to .18. This was indeed an unexpected result because, during the construction of Table P, the authors had noticed a significant number of new FQ coding entries, an occurrence that has been reported in a published paper (Silva & Dias, 2007a).

Table 3 shows iota values for the three FQ variables at the response level and the eight FQ variables at the protocol level. The results of the analyses are rather revealing. At the response level, for FQ— the level of agreement for each answer is high (.81), showing consistency of judgments across tables with regard to perceptual distortion. For FQo it also is still quite high (.72). The lowest level of agreement is found for FQu (.56), a finding justified by the very nature of the variable, which depends more on the subject's own way of looking at the blot or part of it, rather than at the stimulus characteristics of the blot, although the blot features are certainly implicit. Perhaps this point contributes to explaining the lower reliability value for FQu as compared to FQo and FQ—. In this regard, we recall

Ephraim's previously mentioned suggestion that answers coded FQu reveal an assimilating attitude toward the stimulus as opposed to the accommodating attitude revealed by the answers coded FQo.

Results for the protocol-level comparisons are similar. Considering the percentages that control for R, the highest iota value (.85) occurs in the X—%, followed by that of X+% (.72), the lowest (.65) being for Xu%. As for variable XA%, which includes encodings FQ+, FQo, and FQu, the value of the iota is .86, a value that benefits from the dominant number of FQo responses. The iota value for variable WDA% is similarly high, .87, and benefits from removing some of the less dominant FQu, as all of the atypical Dd responses are excluded when computing this variable.

Thus, the value that gets closer to the real differences between FQ encodings using Tables A and P is the iota value of 0.56 obtained through answer-by-answer agreement on the FQu variable. Although no significant differences were found between the means, this variable gives the best account of the encoding differences related to the use of either table. On the other hand, we feel bound to point out the very high answer-by-answer agreement obtained on variables FQo and FQ—, an incontrovertible indicator of the perceptual basis of Rorschach responses.

### DISCUSSION

To a large extent, these results confirm expectations when the perceptual basis of answers to the Rorschach or the determining function of their critical distal bits are recognized and accepted. In this context, it is understandable that the perceptual base is more generally recognized and accepted across cultures and ages insofar as answers coded FQo are concerned. It is as easily recognized and accepted when there is clear incompatibility between the form of the evoked object and the corresponding area of the chosen blot (FQ—). With respect to FQu, this is defined as the form of an object that is easily recognizable but rarely evoked in the answers. These answers seem to require personal or idiosyncratic ways of looking at the blot and are, by and large, responsible for the traditional idea that the number of possible different answers to the Rorschach cards is unlimited. Although this idea is theoretically correct, it is shown to be of negligible practical interest, as evidenced by Exner, who clearly showed how it was difficult to add a new FQo response to his FQ table.

TABLE 2.—Descriptive data and differences in form quality obtained through the use of Exner's American table (A) and a Portuguese table (P) on the protocols of 180 Portuguese boys and girls aged 6 to 10.

Variables	Tables	<i>M</i>	<i>SD</i>	Differences		Cohen's <i>d</i> <sup>a</sup>
				<i>t</i>	Significance (2-Tailed)	
X+%	A	0.39	0.12	-0.21	0.88	-.02
	P	0.39	0.12			
X—%	A	0.31	0.12	-1.44	0.15	-.15
	P	0.33	0.13			
Xu%	A	0.28	0.12	1.68	0.09	.18
	P	0.26	0.12			
XA%	A	0.67	0.12	1.11	0.19	.14
	P	0.66	0.13			
WDA%	A	0.70	0.12	1.31	0.27	.12
	P	0.69	0.13			

<sup>a</sup> Derived from the *Ms* and *SDs*, not from *t*.

TABLE 3.—Values of iota indicating degree of agreement when comparing form quality variables scored according to the Exner and Portuguese tables for codes assigned to single responses (*N* = 4,261) and for summary scores across whole protocols (*N* = 180).

Variables	Codes Assigned to Single Responses	Summary Scores for Whole Protocols
FQo	0.72	0.82
FQ—	0.81	0.93
FQu	0.56	0.87
X+%		0.72
X—%		0.85
Xu%		0.65
WDA%		0.87
XA%		0.86

The important points to note are that in a Rorschach FQ table, FQo registers the lowest number of entries and that the higher and quite close numbers are found in the coded entries on FQ– and FQu (Silva & Dias, 2007a). This indicates, on the one hand, the presence of subjectivity in the answer, be it in a negative perspective or in a merely personal perspective, and, on the other hand, the vast diversity in imagery that the inkblots can generate. The reason for a high *iota* value on the FQ– is explained as a distortion or a perceptual error, an explanation that is widely accepted. However, the reason for the low *iota* value on the FQu rests on the possibility of seeing inkblots or some of their areas as good form objects, that is, objects with forms that are not immediately, easily, and frequently recognized; which, in other words, are not “pregnant” with a prepotent image. Such a possibility of seeing good forms in nonpregnant blot areas is very broad and presumably takes along with it some cultural and educational characteristics, although its meaning, whether personal or subjective, remains important. It is not difficult to recognize that both some cultural aspects and even certain modal interests or situational events might be at the root of some FQu answers across diverse world countries or cultures. It might even happen that certain FQu responses occur in the form of sudden findings, in which the subject sees a good form and acknowledges it with manifestations of happiness, surprise, or both. The unusual feature of these responses—based on the identification of a rare combination of critical bits of the inkblot, coupled with the range of circumstances and motivations of the subject who gives them—might be the reason for the lower reliability of the values for this variable when compared to those of the remaining set of FQ variables. This happens both on the interjudge reliability (see Table 1) and on the results of our comparative work (see Table 3).

During the 1990s, the study of the interjudge reliability in the Comprehensive System gained relevance. In their work, Acklin, McDowell, Verschell, and Chan (2000) suggest kappa was the most adequate statistic to use. During the first few years of the last decade several important works followed: Meyer et al. (2002), Viglione and Taylor (2003), and McGrath et al. (2005). In all of them, the protocol-level and response-level reliability values found for the variable FQu are notably lower than the values for X+% and X–%. The consistency of results across studies suggests that the source of the lower level of interjudge or inter-table reliability for FQu comes from the variable itself, from its very nature and from the criteria that define its encoding that, for instance, allows a particular area of the inkblot to be coded together with responses of different levels of formal organization. Good examples are responses such as “insect with wings” or “map”; in the first case, the object has a specific form, whereas in the second the form is not taken into consideration, nor is it specified.

Results obtained using Table P should be considered, in principle, to have a positive bias, or rather, benefit from being some of the protocols that contributed to the construction of Table P. However, Table 4 presents data that suggest otherwise. When the FQ scores from the full sample of normative data for Portuguese children (Silva & Dias, 2007b), whose protocols were coded with Table A, are compared with FQ scores of our sample of 180 protocols coded with Tables A and P, it seems hardly possible to consider the presence of a positive bias in all FQ variables, because scores in the Table P column are equal to

TABLE 4.—Mean form quality values obtained using Table A in the full normative sample (*N* = 357) and using Tables A and P in the experimental subsample (*N* = 180).

FQ variables	Table A <sup>a</sup>	Table A <sup>b</sup>	Table P <sup>b</sup>
X+%	.38	.39	.39
X–%	.32	.31	.33
Xu%	.30	.28	.26
XA%	.67	.67	.66
WDA%	.69	.70	.69

<sup>a</sup>*N* = 357; total sample of Portuguese children between ages of 6 and 10.  
<sup>b</sup>*N* = 180.

or lower than those in the Table A column (see Table 4). It is difficult to find an objective explanation of this fact.

Nonetheless, among all FQ variables, it is precisely FQu that is responsible for the larger disparity or differences between the encodings of the two tables. This finding leads us back to our initial citation from Exner and Weiner (1995):

Instead, it seems more reasonable to suggest that the items in the Form Quality Table be reviewed for frequency. In other words, if a response currently not found in the Table occurs frequently in a given country or culture, the form quality scoring for that answer should be adjusted accordingly. Likewise, if a response that is currently listed in the Form Quality Table as unusual occurs with a high frequency in the nonpatient records for a country or culture, it should be scored ordinary. This is a very different procedure from that involved in attempting to establish separate normative data. (p. 50)

Taking this suggestion to its logical conclusion one might state, against our initial perspective from our sample of Portuguese children, that the absence of significant differences between the means referred to earlier allows for the inclusion, in its entirety, of normative data from various sources, namely Rorschach international norms.

At this point, a new question arises here concerning our Table P: In view of the results obtained, can the use of this type of table be justified? We believe it is not yet the time to make a decision on this. It would perhaps be wiser to wait for other initiatives of this kind to better judge their relevance. Coming back to our data, it is therefore important to establish the precise meaning of the .56 *iota* value obtained for the FQu variable, one that indicates the number of responses with encodings that do not coincide in both FQ tables considered, and ascertain to what it corresponds. Table 5 presents the distribution of total FQu entries of FQ Table P, showing the number of entries that are also coded FQu in Table A, those being coded as FQo just in Table A, and those that are absent from it, that is, exclusively present in Table P. It reveals two important data points that permit better evaluation of our statements: First, it shows that 330 of the 490 FQu entries (67%) are differently coded in both tables, thus justifying the referred low *iota* value. Second, there are 190 entries (33%) in Table P absent in Table A, a number that seemed to us higher than expected.

We should also offer an explanation for the high .86 *iota* obtained on XA% vis-à-vis the low .56 obtained on FQu. This is due to the much larger number of answers coded FQo as compared to the number of answers on FQu, and also the fact that the FQo is nearly always present in both tables; furthermore,

TABLE 5.—The distribution of Table P FQu entries on all 10 Rorschach cards and their corresponding presence in Table A as FQu, in Table A as FQo, or in Table P only

Cards	I	II	III	IV	V	VI	VII	VIII	IX	X	Total
All Table P FQu	47	52	54	45	28	42	34	55	61	72	490
Present in Table A as FQu	16	11	11	15	14	15	10	24	17	27	160
Present in Table A as FQo	12	14	24	12	7	12	12	11	18	18	140
Present just in Table P	19	27	19	18	7	15	12	20	26	27	190

a considerable number of FQu answers in Table P, 140, are also in Table A.

Yet referring to the pertinence of new FQ tables, whereas Portuguese children norms for FQu between 6 and 10 years of age have mean values from .26 to .33 (Silva & Dias, 2007b), these figures rise to .37 in Portuguese adult norms (Pires, 2007). So, if our data in Table 5 show that 33% of FQu entries are absent from Table A, it seems important to ask whether a similar or higher number might be reached on a Table P for adults. That is why we consider it premature to decide against building new FQ tables.

Finally, given the absence of significant differences between diverse FQ variables, FQu excepted, as a result of using Table A and Table P in coding the same protocols, we feel inclined to consider the chief element responsible for these results to be the perceptual dimension underlying the response process. As a matter of fact, the consistency between the words designating the content and the form requirements of this same content is manifest, be it American or Portuguese. Against all our former expectations, the absence of significant differences between FQ variables is here evidenced. Now we recognize that differences between our means and those published by Exner and Weiner (1995) do not have their source in the use of Exner's Table A for FQ, except for FQu, nor in the use of different languages as suggested by Andronikof-Sanglade (2000). Another origin for those differences must be found. Perhaps the answer lies in their perceptual nature.

### CONCLUSION

Our research shows the effect of the application of two different FQ tables, which in our case affects mainly the FQu variable, when the answer-by-answer method of analysis is used. It also shows the close similarity of FQ assessment of the answers to the Rorschach between Table P (built on 400 protocols of non-patient Portuguese children) and Table A (the North American table built on 9,500 adult protocols of nonpatient, institutionalized, and noninstitutionalized patients; Exner, 2003, p. 122) when applied to the 180 protocols of the Portuguese sample. Such similarity, we underline, attests to the perceptive nature of Rorschach responses. We believe that our results, both for their objectivity and for their novelty, are an important source of clarification and reflection.

### ACKNOWLEDGMENTS

This study was supported by Centro de Investigação Psicológica da Universidade de Lisboa (CIPUL). We would like to express to Gregory J. Meyer our gratefulness for all his valuable

suggestions that have contributed to the better quality of our text.

### REFERENCES

- Acklin, M. W., McDowell, C. J., II, Verschell, M. S., and Chan, D. (2000). Interobserver agreement, intraobserver reliability, and the Rorschach Comprehensive System. *Journal of Personality Assessment*, 74, 15–47.
- Andronikof-Sanglade, A. (2000). Use of the Rorschach Comprehensive System in Europe: State of the art. In R. H. Dana (Ed.), *Cross-cultural and multicultural personality assessment* (pp. 329–344). Mahwah, NJ: Erlbaum.
- Ephraim, D. (2000). Culturally relevant research and practice with the Rorschach Comprehensive System. In R. H. Dana (Ed.), *Cross-cultural and multicultural personality assessment* (pp. 303–327). Mahwah, NJ: Erlbaum.
- Erdberg, P., & Shaffer, T. W. (1999, July). *International symposium on Rorschach nonpatient data: Findings from around the world*. Paper presented at the International Congress of Rorschach and Projective Methods, Amsterdam, The Netherlands.
- Exner, J. E. (1985). *Rorschach workbook for the Comprehensive System* (2nd ed.). Bayville, NY: Rorschach Workshops.
- Exner, J. E. (1990). *Rorschach workbook for the Comprehensive System* (3rd ed.). Asheville, NC: Rorschach Workshops.
- Exner, J. E. (1995). *Rorschach workbook for the Comprehensive System* (4th ed.). Asheville, NC: Rorschach Workshops.
- Exner, J. E. (2001). *Rorschach workbook for the Comprehensive System* (5th ed.). Asheville, NC: Rorschach Workshops.
- Exner, J. E. (2003). *The Rorschach: A Comprehensive System: Vol. 1* (4th ed.). Hoboken, NJ: Wiley.
- Exner, J. E., & Weiner, I. B. (1995). *The Rorschach: A Comprehensive System Vol. 3: Assessment of children and adolescents* (2nd ed.). New York, NY: Wiley.
- Faul, F., Erdfelder, E., Lang, A.-G., & Butcher, A. (2007). G\*Power 3: A flexible statistic power analysis program for the social, behavioural, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Janson, H. (2003). *Rorschach research utilities (Version 1.0.0) user's guide*. Mimeographed.
- McGrath, R. E., Pogge, D. L., Stokes, J. M., Cragnolino, A., Zaccario, M., Hayman, J., . . . , Wayland-Smith, D. (2005). Field reliability of Comprehensive System scoring in an adolescent inpatient sample. *Assessment*, 12, 199–209.
- Meyer, G. J. (2001). Evidence to correct misperceptions about Rorschach norms. *Clinical Psychology: Science and Practice*, 8, 389–396.
- Meyer, G. J., Erdberg, P., & Shaffer, T. W. (2007). Toward international normative reference data for the Comprehensive System. *Journal of Personality Assessment*, 89(Suppl. 1), S201–S216.
- Meyer, G. J., Hilsenroth, M. J., Baxter, D., Exner, J. E., Jr., Fowler, J. C., Piers, C. C., & Resnick, J. (2002). An examination of interrater reliability for scoring the Rorschach Comprehensive System in eight data sets. *Journal of Personality Assessment*, 78, 219–274.
- Meyer, G. J., & Viglione, D. J. (2008). An introduction to Rorschach assessment. In R. P. Archer & S. R. Smith (Eds.), *Personality assessment* (pp. 281–336). New York, NY: Routledge.
- Pires, A. A. (2007). Rorschach Comprehensive System data for a sample of 309 adult nonpatients from Portugal. *Journal of Personality Assessment*, 89(Suppl. 1), S124–S130.

- Shaffer, T. W., Erdberg, P., & Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS-R and MMPI-2. *Journal of Personality Assessment*, 73, 305–316.
- Shaffer, T. W., Erdberg, P., & Meyer, G. J. (2007). Introduction to the JPA Special Supplement on International Samples for the Rorschach Comprehensive System. *Journal of Personality Assessment*, 89(Suppl. 1), S2–S6.
- Silva, D. R., & Dias, A. M. (2007a). A propósito de la Elaboración de una Tabla de Calidad Formal del Rorschach de una muestra de niños portugueses [Are Rorschach normative data presented by Exner valid in extenso for the European population? An essay with a Portuguese sample]. *Revista de la Sociedad Española del Rorschach y Métodos Proyectivos*, 20, 34–38.
- Silva, D. R., & Dias, A. M. (2007b). Rorschach Comprehensive System data for a sample of 357 Portuguese children at five ages. *Journal of Personality Assessment*, 89(Suppl. 1), S131–S141.
- Silva, D. R., Novo, R., & Prazeres, N. (1990). Les Données du Rorschach Présentées par Exner Sont-elles Valables in extenso pour la Population Européenne? Essai avec un Echantillon Portugais [On the creation of a Rorschach Form Quality Table from a Children Portuguese sample]. XIII International Congress of Rorschach and Other Projective Methods, Paris, 1990. *Revista de la Sociedad Española del Rorschach y Otros Métodos Proyectivos*, 3, 33–40.
- Viglione, D. J., & Hilsenroth, M. J. (2001). The Rorschach: Facts, fiction and future. *Psychological Assessment*, 13, 452–471.
- Viglione, D. J., & Taylor, N. (2003). Empirical support for interrater reliability of Rorschach Comprehensive System coding. *Journal of Clinical Psychology*, 59(1), 111–121.
- Wood, J. M., Nezowski, M. T., Garb, H. N., & Lilienfeld, S. O. (2001). The misperception of psychopathology: Problems with the norms of the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice*, 8, 350–373.